DOCUMENT RESUME

ED 256 774 TM 850 015

AUTHOR Lowe, Rosemary H.
TITLE Job Evaluation Technologies and Their Potential for Unbiased Measurement of Job Worth.
PUB DATE Mar 84
NOTE 21p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Bias; Comparative Analysis; *Compensation (Remuneration); *Job Analysis; Latent Trait Theory; Salaries; *Salary Wage Differentials; *Sex Bias; Statistical Analysis
IDENTIFIERS Comparable Worth; *Job Evaluation; Job Value Factors

ABSTRACT

This report examines the process of job evaluation as a method of measuring job worth. Latent trait theory is applied to suggest methods of identifying wage bias. The point factor method of job evaluation selects compensable factors and weights them, using regression analysis. This method is used to maximize the prediction of the existing wage structure for these jobs. The choice and weights of factors reflect the market value of jobs. The a priori method of selecting and weighting compensable factors is preferred by advocates of pay equity. Factors are determined through a process of judgment and negotiation before any jobs are evaluated. Research on psychometric techniques is suggested if job evaluation is to be used to remedy wage discrimination. Several potential sources of bias in job evaluation are suggested. Systematic job evaluations should be conducted more frequently to reduce wage inequities (DWH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED256774

Job Evaluation Technologies and their Potential for Unbiased Measurement of Job Worth

Rosemary H. Lowe

Louisiana State University

Southeastern Psychological Association

March, 1984

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

R. Lowe

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

TM 850 015

In the fall of 1983, Greyhound bus drivers went on strike over a 28% slash in wages, benefits, and anticipated cost-of-living increases which had been proposed by management. That sounded drastic to me. Of course bus drivers would be outraged over that large a cut in their compensation.

However, my feelings changed when I heard that the average compensation package for these bus drivers was $35,000 ("Greyhound Collides," 1983)! I had a Ph.D. and several years work experience struggling with what I considered to be a tough job and bad working conditions. And I was making considerably less than a bus driver!

When we perceive that payment for a job is "unfair," this is a result of a mental evaluation of the job's requirements and a determination that the "worth" of the job is greater than, or less than the payment provided. We are carrying out in our heads the private equivalent of a formal job evaluation.

The issue of comparable worth centers around judgments of pay equity, based on evaluations of the relative worth of jobs. Job evaluation technologies have been used for years to determine a hierarchy of jobs within an organization or an industry, in part for the purpose of pay setting. Some advocates of comparable worth (Gold, 1983, p. 51; Hartmann & Treiman, 1984, p. 63/10-11; Thomsen, 1979) look to job evaluation as a method for establishing pay equity; however, they and others (Blumrosen, 1979; Remick, 1978; Thomsen, 1978, 1981; Treiman, 1979) point out that job evaluation systems themselves may in part account for wage discrimination in favor of jobs predominantly held by men.

This paper will examine the process of job evaluation as a method

of measuring the worth of jobs, focusing on the potential for using job evaluation in a bias-free way. The psychometric perspective of latent trait theory will be applied to suggest ways of identifying bias. Other likely sources of bias will be reviewed, and some suggestions will be made for reducing this systematic error. Throughout, reference will be made to the sparse empirical research on this topic and directions for future research will be suggested.

Job Evaluation Technologies

There are many methods of job evaluation. In a widely cited review, Treiman (1979, p. 4) describes a survey (Akalin, 1970) indicating that the point factor method is the most widely used system in this country. Since the method is relatively well-developed, straightforward, and quantified, some authorities have suggested it is the most amenable to "bias-free implementation" (Gold, 1983, p. 41).

In the point factor method, compensable factors are chosen and weighted, and point values are assigned to different levels on the scale for each factor. A job is then assigned a number of points on each factor, with all jobs being evaluated in terms of the same point maximum for a given factor. Factor weights are applied, and a job's worth is the total of points for all factors.

In a specific application of this method, it is important to ask how the factors and weights are chosen. In some plans, benchmark or key jobs are chosen. They are jobs whose worth and whose wage is consensually agreed upon. Compensable factors are then selected and weighted using regression analysis to maximize the prediction of the existing wage structure for these jobs. Thus the choice and weights of factors reflects the market value of jobs, including the results of

collective bargaining (Milkovich & Cogill, 1984, p. 10/11). This method is called "policy-capturing" because it makes explicit the values or policies underlying an agreed-upon wage structure. Some authors have argued that policy-capturing methods "cannot correct inequities against women's jobs because (their) express purpose is to rationalize an existing structure in which women's jobs are undervalued" (Gold, 1983, p. 53).

The second means of selecting and weighting compensable factors, which is preferred by advocates of pay equity, is the a priori method. In this method, factors are determined through a process of judgment and negotiation, by management with or without workers, before any jobs are evaluated. Although a priori methods have the potential for unbiased measurement of job worth, in the sense that they may result in factors upon which all concerned can agree, it seems clear that bias can be built into these systems also.

For example, the committee report of a job evaluation study conducted by Hay Associates for the State of Illinois states that "the majority...felt that certain management and sex biases are extant in the Hay method" (Commission on the Status of Women, State of Illinois, 1983, June, p. 22). As an illustration, "men's" tools such as jackhammers are rated more specialized than the video display terminals used by many women; "men's" lifting, as in stocking shelves, is rated more strenuous than "women's" lifting of patients.

Job Worth as a Latent Trait

If job evaluation is to be used as a remedy for wage discrimination, researchers and practitioners must pay a great deal more attention to psychometrics. In stark contrast to the large number of

fairly recent studies on test fairness in selection, most empirical research on job evaluation is over 30 years old and thus antedates the EEO legislation of the 1960s.

One popular current psychometric perspective is the latent trait approach, which views a test as an attempt to order persons on some underlying attribute which cannot be directly observed. Thorndike (1982) has described the latent trait approach and the concept of bias as they apply to the measurement of attributes of people. His ideas can be applied to the measurement of attributes of jobs, which is what we mean by job evaluation. Thus, a "test" in Thorndike's terms corresponds to a job evaluation score which attempts to describe the underlying trait of "job worth." Thorndike's "test items" can be understood as factor or subfactor scores which may vary in their reliability, validity, intercorrelation, and item characteristics. One value of a latent trait approach to job evaluation concerns the analysis of bias and the steps which can be taken to eliminate it. Some recommendations adapted from Thorndike will serve to illustrate this perspective, and to suggest directions for empirical research.

---

insert Figure 1 about here

If we could directly measure job worth (which we can't), we could plot the "factor characteristic curve" showing the probabilities of various scores on a factor such as "working conditions," on the scale of job worth. In the absence of a pure measure, we can use total job worth points, salary, or any other independent numerical estimate of worth. Figure 1 shows some such hypothetical curves. Following Thorndike, a

"steep" curve indicates that the score depends almost entirely on the amount of job worth. Thus, curve A represents a factor that is highly correlated with job worth, while curve B illustrates a factor whose score depends on many extraneous things, including unreliability of job factor scores.

Bias in factor scores. In the context of the comparable worth issue, our purpose is to identify bias in the evaluation of men's and women's jobs. At the level of factor scores (Thorndike's "items"), bias would be indicated by differences in the job factor characteristic curves for the same factor when applied to men's and women's jobs. A particular factor should be discarded if it shows a low correlation with job worth score in either men's or women's jobs, or in both groups, or if the difference in correlation for mens' and for women's jobs is large.

Thorndike also suggests dropping items (i.e., job factors) which appear to be inappropriate for one or both groups of jobs, in the view of persons closely acquainted with the jobs being evaluated. This procedure is more relevant to the political than to the statistical unsuitability of factors.

Thus, a psychometric approach leads to the conclusion that the same set of job factors should be used to evaluate women's and men's jobs. This contradicts a common practice of using different job evaluation plans for different categories of workers (e.g., professional vs. clerical). Treiman (1979, pp. 46-48) presents data from the comparable worth study conducted by Norman D. Willis and Associates for the State of Washington. Using universal factors to evaluate all jobs, they found that the monthly dollar return for job evaluation points was directly

related to the proportion of incumbents who were male. That is, universal factor job worth points predicted lower salary levels for women's jobs than for men's; this was not apparent when jobs were evaluated on separate job plans.

When different factors are used for men's and women's jobs, worth is assumed to be a dimension underlying all jobs, yet indices of worth (the job factors) are selected to be different for different job categories. This is like measuring girls on math and boys on reading, and using both scores as comparable indicators of school performance. Of course they are not. Job evaluators must use a set of compensable factors which are equally applicable to jobs of differing content, or the jobs cannot be fairly compared.

The importance of the choice of subfactors is shown in a study by Cascio and O'Reilly (1981), in which subjects were asked to allocate job evaluation points to subfactors, that is, to assign subfactor weights. For three of four factors, both male and female subjects allocated significantly more weight to factors drawn from a "male" evaluation plan than to the "female" subfactors, even though the factors were not labeled as such. These results support the argument that jobs may receive more or fewer job worth points as a function simply of the choice of factors or subfactor definitions. Treiman (1979, p. 38) also presents comparative data from two actual job evaluation projects showing dramatic differences in the evaluated worth of some jobs as a function of the evaluation system used.

Bias in job worth scores. Thorndike describes three ways to identify bias in "complete tests," which corresponds in our analysis to the composite job worth score. A difference in mean job worth score

between male and female jobs may indicate bias resulting from choice of factors which are more relevant to one category of jobs than to the other. One appropriate test for such bias would be to apply to female jobs a set of compensable factors specifically designed for male jobs, and vice versa. "If group differences are eliminated or reversed by (factors) specifically designed to favor the previously low-scoring group, the hypothesis of bias in (factor) content receives support. If not, the hypothesis is largely demolished" (Thorndike, 1982, p. 236). This test could be carried out using the male and female subfactors suggested in an article by Remick (1978).

A second way to identify job worth bias is to examine subgroup differences in reliability of scores. Though such data may exist in proprietary job evaluation studies, they have not been reported recently in the literature on job evaluation. One implication of lower reliability of job worth scores for one group is that the correlation of job worth with a wage criterion will be lower as well.

A third sign of bias, and the one analogous to the focus of the test fairness issue, is the relationship of job worth scores to a criterion such as wages or salary. The Cleary model for test fairness, which is also used in the Uniform Guidelines (1978), specifies a comparison of the slopes and intercepts of the regression of predictor (here, job worth) on criterion (wages) for the two subgroups. "Consistent non-zero errors of prediction" (Cleary, 1968, p. 115) will be made when the slopes and/or intercepts are significantly different. Evidence from actual job evaluation studies suggests that, by this standard, bias exists.

---

insert Figure 2 about here

Figure 2 presents data from the State of Washington study conducted by Norman D. Willis and Associates in 1974, adapted by Treiman (1979). Job worth points are plotted on the horizontal axis, and average monthly salary on the vertical. Visual inspection of the lower regression line (for predominantly female jobs) suggests that its slope is different than that for the predominantly male jobs plotted on the upper regression line.

More recent data analyzed by Hay Associates for the State of Illinois suggest another point. In Figure 3, regressions of job worth points on monthly dollars are presented for female-dominated job classes (lower line) and for male-dominated job classes (upper straight line). However, the variability about the male line was such that the best fit for males was actually a curvilinear regression, as indicated by the upper, bent line. A significant curvilinear term in the regression for males, but not for females, means that the slopes are different and thus, following Thorndike, bias exists. The committee which prepared the narrative report interpreted the curvilinear results to mean that factors other than job content affect salary for male-dominated classes, and that the male labor market is more segmented than that for females (CSW, 1983, June).

---

insert Figure 3 about here

Likely sources of bias in job evaluation

Systematic underestimation of the worth of women's jobs, or overestimation of the worth of men's, can occur at several steps in the job evaluation process. Several writers (Blumrosen, 1979; Milkovich & Broderick, 1982; Remick, 1978; Treiman, 1979) have discussed points at which bias is especially likely:

1. Job analysis may omit or overemphasize information on certain duties, and the difficulty of duties may be inaccurately estimated.
2. The written job description or specification can be no more accurate or complete than the job analysis, and in fact may be less so due to omissions or distortions.
3. The selection, definition, and scaling of compensable factors may be biased. For example, Treiman (1979, p. 32) points out that "effort" is usually defined in terms of strength rather than fatigue levels, and "manual skill" in terms of heavy tools rather than manual dexterity. He cites one evaluation plan which rated "typist" as requiring one month of training time, while "truck driver" was said to require 12 months! With such scoring, "women's" jobs will often be undervalued.
4. Compensable factors may be weighted to the disadvantage of women's jobs. For example, a widely-used industrial plan weights "job experience" more heavily than "formal education," a procedure which is likely to result in undervaluation of jobs held mainly by women (Treiman, 1979, p. 10). When weights are chosen for their ability to predict wage rates, and especially when benchmark jobs are predominantly male, this is especially

likely. Furthermore, as Schwab (1980) points out, the apparent factor weights will be misleading when the variances of scores on factors are unequal; factors on which scores are quite variable will be "weighted" more heavily, regardless of the explicit weights chosen.

5. Different evaluation plans may be used for job categories within the same firm; this renders job worth scores noncomparable across job categories.
6. Validation of evaluation plans by an external wage structure which is itself discriminatory will build inequities into the system.
7. Wage surveys may themselves be hastily done, unreliable, invalid, or the equivalent of price-fixing.

Some partial remedies for job evaluation bias have been suggested. At a very minimum, professionals should adhere to standards of good practice in job analysis and description, including review of documents by incumbents (Thompson & Thompson, 1982). Compensable factors, definitions, and weights can be developed by committees with diverse representation from among employees; workers can have access to information about the final plan and mechanisms can be developed for career movement to jobs of higher worth, when appropriate.

Research is needed on the psychometric properties of various evaluation methods and processes. For example, Schwab and Wichern (1983) have recently shown logically that under certain assumptions and conditions, undervaluation of predominantly female jobs may actually lead to <u>overpayment</u> of some women relative to their "true wage." However, bias in the wage criterion does have the generally predicted

effect of depressing wages for females in nonkey jobs. These predictions depend on the relative proportions of male and female jobs, their distribution on the wage scale, and whether predictions are made from salaries for male jobs, female jobs, or both. Their predictions should be submitted to empirical test with actual job evaluation data.

Several authors (Milkovich & Broderick, 1982; Schwab, 1980; Schwab & Wichern, 1983; Treiman, 1979; Treiman & Hartmann, 1981) have called for experimental research on the judgment processes involved in job evaluation, such as choice and weighting of factors and the assignment of factor points. Very little published research, for example, deals with the effect of incumbent characteristics upon evaluations of jobs.

Another empirical question concerns the role of supply and demand in wage-setting. Many writers (Forbes, 1984; Seligman, 1984) credulously assume that market wages closely reflect supply and demand, which it can be argued they do not (e.g., Treiman & Hartmann, 1981). If different pay for comparable jobs does result from forces of supply and demand, observed wages should be more accurately predicted from job evaluation points when equations include a term describing shortages or excesses of qualified personnel (Wallace, 1983). Failure to confirm this prediction would support the notion of artificial restrictions on "natural" wage increases.

Finally, systematic job evaluations should be done more often. Cooper and Barrett (1984), in their review of court cases concerning compensation management, show that competently conducted and up-to-date job evaluations have been useful evidence for both plaintiff and defendant in this litigation, but that job evaluations are seldom introduced. It may not be possible for job evaluation to be completely

unbiased, but the procedure can be less biased than the alternatives. Recent litigation relevant to comparable worth (such as Gunther and the State of Washington case) has arisen from a failure to implement pay equity recommendations based on job evaluation studies-- and not from the question of fairness in the job evaluation itself. (For a review of recent comparable worth litigation, including introduction of job evaluations as evidence, see Golper, 1983.)

In closing, I'd like to suggest a good research topic for someone interested in pay equity and job evaluation. Last year during the professional football strike, an article appeared in which the work of the Dallas Cowboy Cheerleaders was described. I was surprised to learn that they are paid a grand sum of $15.00 each per game. Management said the women would never go on strike, since there were so many dying to take their places! Now there's an opportunity for a study.

References

Akalin, M. T. (1970). Office job evaluation. Des Plaines, IL: Industrial Management Society.

Blumrosen, R. G. (1979). Wage discrimination, job segregation, and women workers. Women's Rights Law Reporter, 6, 19-57.

Caccio, W. F., & O'Reilly, C. A., III (1981, August). Comparable worth and job evaluation: The biasing effects of subfactors and contextual cues. Paper presented at the annual convention of the American Psychological Association, Los Angeles.

Cleary, T. A. (1968). Test bias: Prediction of grades of negro and white students in integrated colleges. Journal of Educational Measurement, 5, 115-124.

Commission on the Status of Women, State of Illinois. (1983). A study of job classifications currently used by the State of Illinois to determine if sex discrimination exists in the classification system. Springfield, IL: Author.

Cooper, E. A., & Barrett, G. V. (1984). Equal pay and gender: Implications of court cases for personnel practices. Academy of Management Review, 9, 84-94.

Forbes, M. S., Jr. (1984). Fact and comment II: A pernicious, destructive idea. Forbes, 133, 23.

Gold, M. E. (1983). A dialogue on comparable worth. Ithaca, NY: ILR Press, Cornell University.

Golper, J. B. (1983). The current legal status of "Comparable Worth" in the federal courts. Labor Law Journal, 34, 563-580.

Greyhound collides head on with its union. (1983, November 21). Business Week, p. 47.

Hartmann, H. I., & Treiman, D. J. (1984). Notes on the NAS Study of equal pay for jobs of equal value. In M. L. Rock (Ed.), Handbook of wage and salary administration (2nd ed., pp. 63/1 - 63/16). New York: McGraw-Hill.

Milkovich, G. T., & Broderick, R. (1982). Pay discrimination: Legal issues and implications for research. Industrial Relations, 21, 309-317.

Milkovich, G. T., & Cogill, C. J. (1984). Measurement as an issue in analysis and evaluation of jobs. In M. L. Rock (Ed.), Handbook of wage and salary administration (2nd ed., pp. 10/1 - 10/14). New York: McGraw-Hill.

Remick, H. (1978). Strategies for creating sound, bias-free job evaluation plans. In Job evaluation and EEO: The emerging issues. New York: Industrial Relations Counselors, Inc.

Schwab, D. P. (1980). Job evaluation and pay setting: Concepts and practices. In E. R. Livernash (Ed.), Comparable worth: Issues and alternatives (pp. 51-77). Washington, DC: Equal Employment Advisory Council.

Schwab, D., & Wichern, D. W. (1983). Systematic bias in job evaluation and market wages: Implications for the Comparable Worth debate. Journal of Applied Psychology, 68, 60-69.

Seligman, D. (1984). Economics undone, or Don Juan meets Dumbo. Fortune, 109, 31-32.

Thompson, D., & Thompson, T. (1982). Court standards for job analysis in test validation. Personnel Psychology, 35, 865-874.

Thomsen, D. J. (1978). Eliminating pay discrimination caused by job evaluation. Personnel, 55, 11-22.

Thomsen, D. J. (1979). Considerations for non-biased systems: Principles for design and audit of job evaluation plans. In J. A. Grune (Ed.), Manual on pay equity: Raising wages for women's work (pp. 109-112). Washington, DC: Committee on Pay Equity.

Thomsen, D. J. (1981). Compensation and benefits. Personnel Journal, 60, 348-354.

Thorndike, R. L. (1982). Applied psychometrics. Boston: Houghton Mifflin Co.

Treiman, D. J. (1979). Job evaluation: An analytic review. Washington, DC: National Academy of Sciences.

Treiman, D. J., & Hartmann, H. I., (Eds.) (1981). Women, work, and wages: Equal pay for jobs of equal value. Washington, DC: National Academy Press.

Uniform Guidelines on Employee Selection (1978). Federal Register, 43, 38295-38309.

Wallace, M. J., Jr. (1983). Methodology, research practice, and progress in personnel and industrial relations. Academy of Management Review, 8, 6-13.

Figure 1

Factor Characteristic Curves

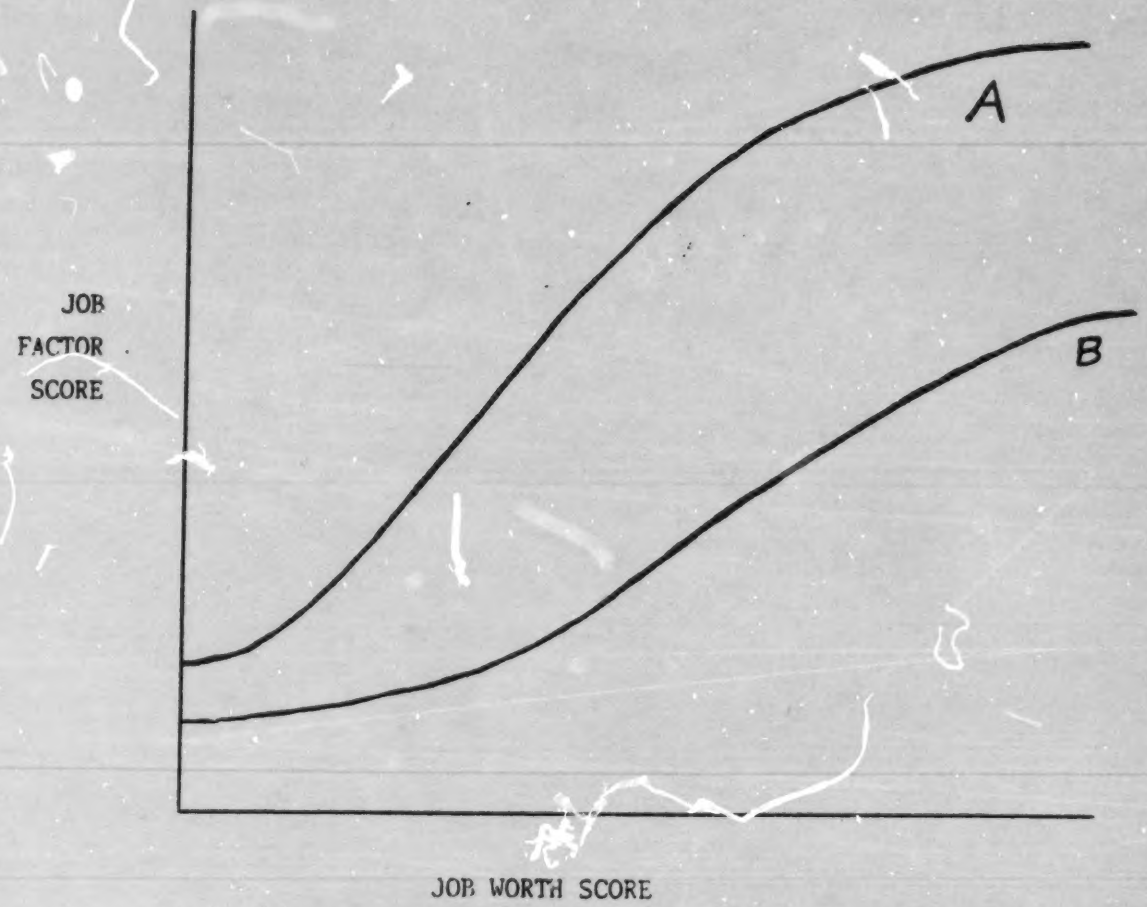


Figure 2
State of Washington Data

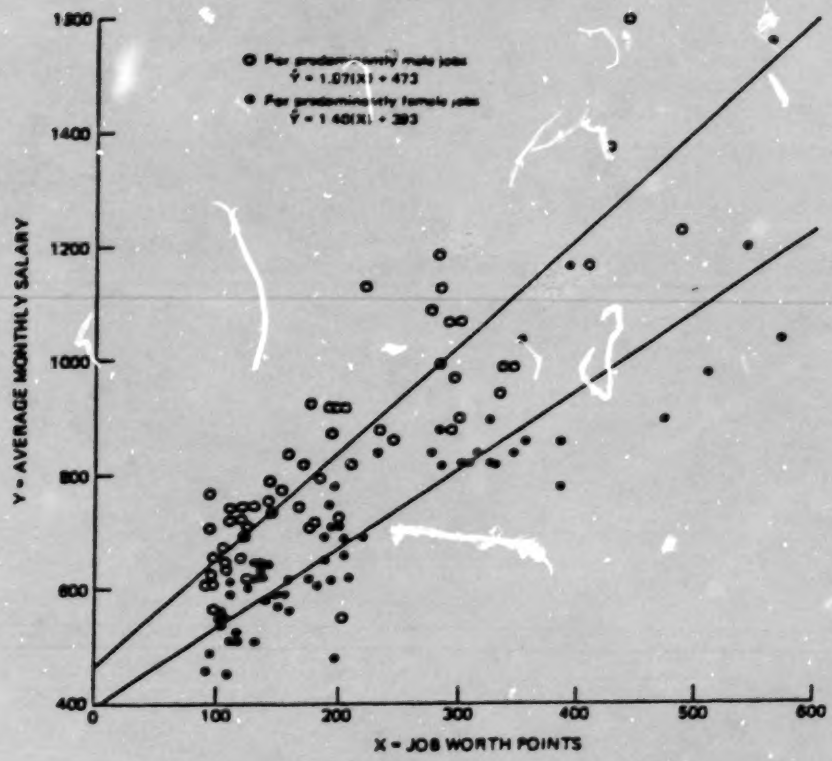


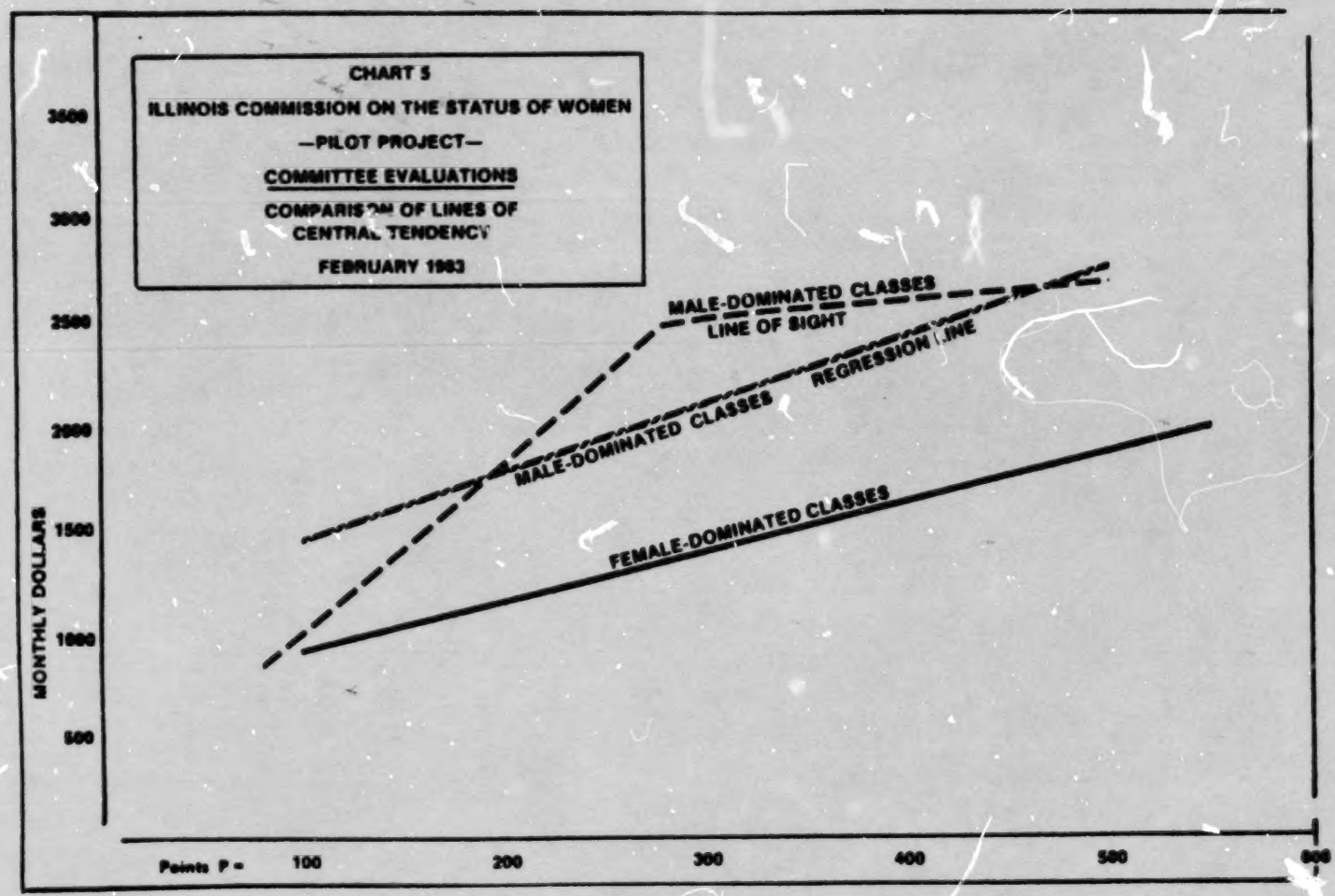
Figure 2. Scatterplot of monthly salaries by job worth points, for 59 predominantly male jobs and 62 predominantly female jobs in the Washington State public service.

SOURCE: Norman D. Willis and Associates (1974) State of Washington Comparable Worth Study. Seattle: Norman D. Willis and Associates. Reprinted by permission.

BEST COPY AVAILABLE

Figure 3

State of Illinois Pilot Study



BEST COPY AVAILABLE